TS

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/818,581	03/28/2001	Yasuo Okutani	862.C2175	3650

5514 7590 12/10/2003

FITZPATRICK CELLA HARPER & SCINTO
30 ROCKEFELLER PLAZA
NEW YORK, NY 10112

EXAMINER	
TRAN, VINCENT V	

ART UNIT	PAPER NUMBER
2655	7

DATE MAILED: 12/10/2003

Please find below and/or attached an Office communication concerning this application or proceeding.

TS

Office Action Summary	Application No.	Applicant(s)	
	09/818,581	OKUTANI ET AL.	
	Examiner	Art Unit	
	vincent v tran	2655	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
- Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☐ Responsive to communication(s) filed on *28 March 2001*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☐ Claim(s) _____ is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-7,10-17,20 and 21* is/are rejected.

7)☒ Claim(s) *8,9,18 and 19* is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *28 March 2001* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. §§ 119 and 120

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

13)☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. § 119(e) (to a provisional application) since a specific reference was included in the first sentence of the specification or in an Application Data Sheet. 37 CFR 1.78.

a) ☐ The translation of the foreign language provisional application has been received.

14)☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. §§ 120 and/or 121 since a specific reference was included in the first sentence of the specification or in an Application Data Sheet. 37 CFR 1.78.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO-1449) Paper No(s) 5 .
4) ☐ Interview Summary (PTO-413) Paper No(s). _____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: .

DETAILED ACTION

Allowable Subject Matter

1. Claims 8-9 and 18-19 are objected to as being dependent upon a rejected base claim, but would be allowable if rewritten in independent form including all of the limitations of the base claim and any intervening claims.

The following is an examiner's statement of reasons for allowability:

Regarding claims 8-9 and 18-19, the Huang et al. reference discloses the method or the apparatus, wherein the distortion output is obtained on the basis of a concatenation distortion produced upon concatenating the synthesis unit to another synthesis unit (col.1, ln.66 – col.2, ln.2), and a modification distortion produced upon modifying the synthesis unit (col.7, ln.26-34 and ln.44-48 and Fig.7, elements #174, #176, #178, #180 and #182).

This reference does not specifically teach nor fairly suggest a method or an apparatus, which has a table that stores the modification distortion/concatenation distortion, and determines the modification distortion/concatenation distortion by looking up the table.

Claim Rejections - 35 USC § 102

2. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

3. Claims 1-3, 10-13 and 20-21 are rejected under 35 U.S.C. 102(b) as being anticipated by Huang et al. (U.S. Patent No. 5,913,193).

Referring to claim 1, Huang et al. disclose a speech synthesis apparatus comprising:

distortion output means for obtaining a distortion (col.9, ln.28-34) produced upon modifying a synthesis unit (instance, col.4, ln.48-52) on the basis of predetermined prosody information (prosody engine, Fig.1, element #35 and col.4, ln.33-42; Fig.7, element #174); and

unit registration (selection, Fig.1, element #23) means for selecting the synthesis unit to be registered in a synthesis unit inventory (speech synthesizer, element #36 and col.4, ln.48-52) used in speech synthesis on the basis of the distortion output from the distortion output means.

Referring to claim 2, Huang et al. disclose the apparatus, wherein the distortion output means obtains the distortion on the basis of a concatenation distortion produced upon concatenating the synthesis unit to another synthesis unit (col.1, ln.66 – col.2, ln.2), and a modification distortion produced upon modifying the synthesis unit (col.9, ln.53-56; col.7, ln.26-34 and ln.44-48 and Fig.7, elements #174, #176, #178, #180 and #182).

Referring to claim 3, Huang et al. disclose the apparatus further comprising:

text input means for inputting text data (col.3, ln.16);

language analysis (NLP, natural language processor) means for performing language analysis of the input text data (col.4, ln.23-26); and

prosody generation (prosody engine) means for generating the predetermined prosody information on the basis of an analysis result of the language analysis means (col.4, ln.25-26 and ln.38-46).

Referring to claim 10, Huang et al. disclose the apparatus further comprising speech synthesis means for producing synthetic speech of text data using the synthesis unit inventory (speech synthesizer, Fig.1, element #36 and col.4, ln.48-52).

Referring to claim 11, Huang et al. disclose a speech synthesis method comprising:

distortion output step of obtaining a distortion (col.9, ln.28-34) produced upon modifying a synthesis unit (instance, col.4, ln.48-52) on the basis of predetermined prosody information (prosody engine, Fig.1, element #35 and col.4, ln.33-42; Fig.7, element #174); and

unit registration (selection, Fig.1, element #23) step of selecting the synthesis unit to be registered in a synthesis unit inventory (speech synthesizer, element #36 and col.4, ln.48-52) used in speech synthesis on the basis of the distortion output from the distortion output step.

Referring to claim 12, Huang et al. disclose the method, wherein the distortion output step, the distortion is obtained on the basis of a concatenation distortion

produced upon concatenating the synthesis unit to another synthesis unit (col.1, ln.66 – col.2, ln.2), and a modification distortion produced upon modifying the synthesis unit (col.9, ln.53-56; col.7, ln.26-34 and ln.44-48 and Fig.7, elements #174, #176, #178, #180 and #182).

Referring to claim 13, Huang et al. disclose the method further comprising:

inputting text data (col.3, ln.16);

performing language analysis (NLP, natural language processor) of the input text data (col.4, ln.23-26); and

generating the predetermined prosody information on the basis of an analysis result of the language analysis step (prosody engine; col.4, ln.25-26 and ln.38-46).

Referring to claim 20, Huang et al. disclose the method further comprising the synthetic speech of text data using the synthesis unit inventory (speech synthesizer, Fig.1, element #36 and col.4, ln.48-52).

Referring to claim 21, Huang et al. disclose a computer readable storage medium storing a program that implements a speech synthesis method (col.10, ln.15).

Claim Rejections - 35 USC § 103

4. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and

the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

5. Claims 4-7 and 14-17 are rejected under 35 U.S.C. 103(a) as being unpatentable over Huang et al. in view of Campbell et al. (U.S. Patent No. 6,366,883).

Referring to claim 4, Huang et al. disclose the apparatus comprising:

selecting the best synthesis unit sequence with reference to the distortion determined based on the concatenation and modification distortions (col.9, ln.44-48 and Fig.7, elements #180 and #182); and

wherein the unit registration (selection, Fig.1, element #23) means selects a synthesis unit to be registered in the synthesis unit inventory (speech synthesizer, element #36 and col.4, ln.48-52).

Huang et al. do not specifically disclose an apparatus for obtaining Nbest sequences of a synthesis unit sequence.

However, Campbell et al. teach an apparatus for obtaining Nbest sequences (N1 best phoneme) of a synthesis unit sequence (col.16, ln.55-60; Fig.5, elements S24, S26, S27 and S28). The advantage of using the teaching of Campbell et al. in Huang et al. would have been to allow the speech synthesis apparatus to search for a combination of phoneme candidates that minimizes the cost (col.3, ln.19-20).

Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention to modify the apparatus of Huang et al., to obtain Nbest sequences of a synthesis unit sequence, as taught by Campbell et al., in order to better

obtain a voice quality closer to the natural voice by selecting the best sequence of sound unit.

Referring to claim 5, Huang et al. disclose the apparatus for selecting the best unit sequence based on minimum accumulated distortion (col.9, ln.44-48).

Huang et al. do not specifically disclose an apparatus for selecting a synthesis unit to be register in the synthesis unit inventory on the basis of a weighted sum of the concatenation and modification distortion.

However, Campbell et al. teach an apparatus for selecting the synthesis unit to be registered in the synthesis unit inventory on the basis of a weighted sum of the concatenation and modification distortion (col.2, ln.37-49 and col.12, ln.1-14). The advantage of using the teaching of Campbell et al. in Huang et al. would have been to allow the speech synthesis apparatus to minimize the target cost and concatenation cost.

Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention to modify the apparatus of Huang et al. which selects the synthesis unit to be registered in the synthesis unit inventory by weighting the sum of the distortions, as taught by Campbell et al., in order to better obtain a voice quality closer to the natural voice by weighting the distortion according to their audibility (col.1, ln.66).

Referring to claim 6, Huang et al. disclose the apparatus wherein the distortion output means determines the modification distortion using an Euclidean distance between synthesis units (col.9, ln.12-15).

Huang et al. do not specifically disclose an apparatus for determining the concatenation distortion using an Euclidean cepstral distance between synthesis units.

However, Campbell et al. teach an apparatus for determining the concatenation distortion using an Euclidean cepstral distance between synthesis units (col.12, ln.8-9; col.16, ln.44-46 and Fig.5, element S23). The advantage of using the teaching of Campbell et al. in Huang et al. would have been to allow the speech synthesis apparatus to minimize the connection cost between speech units (col.13, ln.8-9).

Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention to modify the apparatus of Huang et al. by substituting an Euclidean cepstral distance for an Euclidean distance, as taught by Campbell et al., in order to enhance sound quality by selecting the closest phoneme that are adjacent to one another according the Euclidean cepstral distance.

Referring to claim 7, Huang et al. disclose the apparatus wherein the distortion output means determines the modification distortion using an Euclidean distance between synthesis units before (col.9, ln.7-15) and after modification (col.9, ln.15-22).

Huang et al. do not specifically disclose an apparatus for determining the concatenation distortion using an Euclidean cepstral distance between synthesis units.

However, Campbell et al. teach an apparatus for determining the concatenation distortion using Euclidean cepstral distance between synthesis units (col.12, ln.8-9;

col.16, ln.44-46 and Fig.5, element S23). The advantage of using the teaching of Campbell et al. in Huang et al. would have been to allow the speech synthesis apparatus to minimize the connection cost between phoneme pieces (col.13, ln.8-9).

Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention to modify the apparatus of Huang et al. by substituting an Euclidean cepstral distance for an Euclidean distance, as taught by Campbell et al., in order to enhance sound quality by selecting the closest phoneme that are adjacent to one another according the Euclidean cepstral distance.

Referring to claim 14, Huang et al. disclose the method comprising:

selecting a best synthesis unit sequence with reference to the distortion determined based on the concatenation and modification distortions (col.9, ln.46-48 and Fig.7, element #182); and

wherein the unit registration (selection, Fig.1, element #23) means selects a synthesis unit to be registered in the synthesis unit inventory (speech synthesizer, element #36 and col.4, ln.48-52).

Huang et al. do not specifically disclose a method for obtaining Nbest sequences of a synthesis unit sequence.

However, Campbell et al. teach a method for obtaining Nbest sequences (N1 best phoneme) of a synthesis unit sequence (col.16, ln.55-60; Fig.5, elements S24, S26, S27 and S28). The advantage of using the teaching of Campbell et al. in Huang et al. would have been to allow the method to search for a combination of phoneme candidates that minimizes the cost (col.3, ln.19-20).

Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention to modify the method of Huang et al., to obtain Nbest sequences of a synthesis unit sequence, as taught by Campbell et al., in order to better obtain a voice quality closer to the natural voice by selecting the best sequence of sound unit.

Referring to claim 15, Huang et al. disclose the method for selecting the best unit sequence based on minimum accumulated distortion (col.9, ln.46-48).

Huang et al. do not specifically disclose a method for selecting the synthesis unit to be register in the synthesis unit inventory on the basis of a weighted sum of the concatenation and modification distortion.

However, Campbell et al. teach a method for selecting the synthesis unit to be registered in the synthesis unit inventory on the basis of a weighted sum of the concatenation and modification distortion (col.2, ln.37-49 and col.12, ln.1-14). The advantage of using the teaching of Campbell et al. in Huang et al. would have been to allow the method to minimize the target cost and concatenation cost.

Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention to modify the method of Huang et al. which selects the synthesis unit to be registered in the synthesis unit inventory by weighting the sum of the distortions, as taught by Campbell et al., in order to better obtain a voice quality closer to the natural voice by weighting the distortion according to their audibility (col.1, ln.66).

Referring to claim 16, Huang et al. disclose the method wherein the distortion output step, the concatenation distortion is determined by using Euclidean distance between synthesis units (col.9, ln.12-15).

Huang et al. do not specifically disclose a method for determining the concatenation distortion using a Euclidean cepstral distance between synthesis units.

However, Campbell et al. teach a method for determining the concatenation distortion using Euclidean cepstral distance between synthesis units (col.12, ln.8-9; col.16, ln.44-46 and Fig.5, element S23). The advantage of using the teaching of Campbell et al. in Huang et al. would have been to allow the speech synthesis apparatus to minimize the connection cost between phoneme pieces (col.13, ln.8-9).

Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention to modify the method of Huang et al. by substituting an Euclidean cepstral distance for an Euclidean distance, as taught by Campbell et al., in order to enhance sound quality by selecting the closest phoneme that are adjacent to one another according the Euclidean cepstral distance.

Referring to claim 17, the method wherein the distortion output step, the modification distortion is determined by using an Euclidean distance between synthesis units before (col.9, ln.7-15) and after modification (col.9, ln.15-22).

Huang et al. do not specifically disclose a method for determining the concatenation distortion using Euclidean cepstral distance between synthesis units.

However, Campbell et al. teach a method for determining the concatenation distortion using an Euclidean cepstral distance between synthesis units (col.12, ln.8-9;

col.16, ln.44-46 and Fig.5, element S23). The advantage of using the teaching of Campbell et al. in Huang et al. would have been to allow the speech synthesis apparatus to minimize the connection cost between phoneme pieces (col.13, ln.8-9).

Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention to modify the method of Huang et al. by substituting an Euclidean cepstral distance for an Euclidean distance, as taught by Campbell et al., in order to enhance sound quality by selecting the closest phoneme that are adjacent to one another according the Euclidean cepstral distance.

Conclusion

6. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure Kondo et al. (U.S. Patent No. 6,405,169) teach a speech synthesis apparatus which can produce synthetic speech of a high quality with reduces distortion.

7. Any inquiry concerning this communication or earlier communications from the examiner should be directed to the examiner **Vincent V. Tran** whose E-mail address:

Vincent.tran@USPTO.GOV.

Phone number: (703) 305-1817

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Mr. Talivaldis Ivars Smits, can be reached on (703) 306-3011.

Any inquiry of a general natural or relating to the status of this application should be directed to the Technology Center 2600 receptionist whose telephone number is (703) 305-4700.

8. Any response to this action should be mailed to:

Commissioner of Patents and Trademarks

P.O. Box 1450

Alexandria, VA 22313-1450

Or faxed to:

(703) 872-9314

Hand-delivered responses should be brought to Crystal Park II, 2121 Crystal Dr, Arlington VA, Sixth Floor (Receptionist, Tel. No. 703-305-4700).

Art Unit 2655

VINCENT V. TRAN

Date: November 19, 2003

TALIVALDIS IVARS SMITS
PRIMARY EXAMINER